Molecular Dynamics of Cryptophane and Its Complexes with Tetramethylammonium and Neopentane Using a Continuum Solvent Model

MICHAEL J. POTTER, PAUL D. KIRCHHOFF,*
HEATHER A. CARLSON, J. ANDREW MCCAMMON

Department of Chemistry and Biochemistry and Department of Pharmacology, University of California at San Diego, La Jolla, CA 92093-0365

Received 30 June 1998; accepted 8 February 1999

ABSTRACT: Time scales currently obtainable in explicit–solvent molecular dynamics simulations are inadequate for the study of many biologically important processes. This has led to increased interest in the use of continuum solvent models. For such models to be used effectively, it is important that their behavior relative to explicit simulation be clearly understood. Accordingly, 5 ns stochastic dynamics simulations of a derivative of cryptophane-E alone, and complexed with tetramethylammonium and neopentane were carried out. Solvation electrostatics were accounted for via solutions to the Poisson equation. Nonelectrostatic aspects of solvation were incorporated using a surface area-dependent energy term. Comparison of the trajectories to those from previously reported 25 ns explicit–solvent simulations shows that use of a continuum solvent model results in enhanced sampling. Use of the continuum solvent model also results in a considerable increase in computational efficiency. The continuum solvent model is found to predict qualitative structural characteristics that are similar to those observed in explicit solvent. However, some differences are significant, and optimization of the continuum parameterization will be required for this method to become a efficient alternative to explicit–solvent

* *Present address:* Molecular Simulations, Inc., 9685 Scranton Rd., San Diego, CA 92121-3752

Correspondence to: M. J. Potter, Center for Advanced Research in Biotechnology, National Institute of Standards and Technology, 9600 Gudelsky Drive, Rockville, MD, 20850-3479

Journal of Computational Chemistry, Vol. 20, No. 9, 956–970 (1999)

CCC 0192-8651 / 99 / 090956-15

simulation. © 1999 John Wiley & Sons, Inc. J Comput Chem 20: 956–970, 1999

Keywords: molecular dynamics; implicit solvent; Poisson–Boltzmann; cryptophane; continuum solvent

Introduction

Large-scale conformational change plays a central role in a vast array of biologically important processes, including inter- and intracellular signaling, regulation, and transport. These responses are often ''triggered'' by environmental perturbations such as the binding and release of charged ligands, the creation or destruction of charge via protonation or reaction, protein–protein association, and lipid–membrane association. Computer simulation has the potential to greatly aid our understanding of these systems. Knowledge gained from simulation could be used to help design therapeutics for a wide variety of disorders. Unfortunately, the large size of the systems and the time scale of such responses often renders conventional explicit–solvent simulation techniques impractical.

A large fraction of the computational effort involved in an explicit–solvent molecular dynamics simulation is associated with the many thousands of solvent molecules required to adequately represent a solution environment. The detailed behavior of the vast majority of these molecules is often not of interest. This raises the possibility of increased computational efficiency through the use of an implicit solvent model. A particularly promising approach is the replacement of solvent molecules with a dielectric continuum.[1,2] The electrostatic portion of the solvation energy can be obtained via solution of the Poisson–Boltzmann equation. A surface area-dependent term is added to account for nonelectrostatic effects. This model has been very successful in the calculation of solvation energies in comparison to both experiment and explicit–solvent simulations.[3–5] Recent developments concerning the calculation of electrostatic solvation forces using this model[6–8] have led to its application in molecular dynamics simulations of a number of systems.[7,9–13]

Because continuum electrostatic solvent models neglect the viscous drag and random buffeting exerted by the solvent on the solute, it is necessary to incorporate them through the use of a stochastic dynamics technique.[14] This approach also allows acceleration of conformational sampling via artificial reduction of solvent damping effects.[15,16]

To fully take advantage of the Poisson–Boltzmann, stochastic dynamics model (PBSD), it is important that its behavior relative to explicit simulation be clearly defined. This is especially important, as most current force fields were developed for explicit–solvent simulation, not for use with continuum electrostatics calculations. Those that were optimized for continuum calculations were designed to reproduce solvation energies, and it is not certain that they provide an accurate description of inter- and intramolecular interactions. The use of the PBSD model will also lead to differences in the dynamics of the system. Average conformational properties should be well reproduced, but short time dynamics are likely to be inaccurate relative to explicit simulation. In fact, it is this difference in kinetic behavior, without alteration of equilibrium-average properties, that is one of the motivations for use of the model. It has been shown that the rate of conformational transition in small systems can be increased through reduction in the friction constant used in the stochastic simulation.[16] This is likely to be particularly advantageous for transitions that require solvent displacement. Such transitions are extremely important (e.g., during ligand binding, hinge bending of protein domains, etc.), and they are often very difficult, if not impossible, to treat in explicit simulations.

The effectively instantaneous relaxation of solvent obtained with an equilibrium continuum model also facilitates conformational change. By maximally stabilizing small fluctuations that, in explicit simulation, would occur on a time scale that would not permit full solvent relaxation, kinetic barriers to conformational change are lowered. This property is at once a benefit and drawback of the method. If one is interested in efficient conformational sampling or incorporating changes in ionic strength, it can be very useful. However, if a single, stable structure is being studied, the in-

crease in fluctuations can allow significantly more conformational flexibility than may be desired.

Therefore, a controlled comparison of PBSD to explicit–solvent simulation would be very useful. This comparison would allow differences in conformational sampling and structural stability to be assessed, and provide information as to the applicability of explicit–solvent force fields to continuum–solvent simulations. Recently reported explicit–solvent simulations of a cryptophane system[17–19] provide an excellent opportunity for such a comparison. Cryptophanes are an interesting group of cage-like molecules that are known to bind a wide variety of guests.[20–22]

A schematic diagram of the cryptophane used in this study is given in Figure 1. The molecule is a hexaacid derivative of cryptophane-E. It contains two cyclotryveratrylene groups, each composed of three aromatic rings connected by methylene bridges. Each aromatic ring is bonded to an acetic acid group via an ether linkage. The two cyclotryveratrylene groups are joined together by three phenolic propyl linkers to form a roughly spherical molecule with a central cavity. Propyl linkers, along with the acetic acid groups, form pores through which a guest must pass to bind within the cryptophane.

This report presents a detailed comparison of 5 ns PBSD simulations of the cryptophane alone and its complexes with tetramethylammonium (TMA) and neopentane (NEO) to previously published results of 25 ns explicit–solvent simulations.

These cryptophane systems are small enough to permit thorough simulation, yet large enough to exhibit behavior more representative of that expected in protein systems than previously studied small molecules.[10–13] Cryptophane's ability to bind guests very similar in structure, but with different net charge also makes it an appealing target for this study.

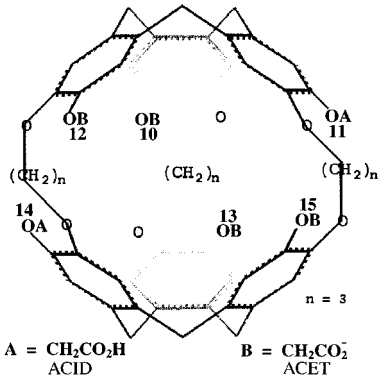

FIGURE 1. The structure of the hexaacid derivative of cryptophane-E used in this study. ACID groups are labeled A and ACET groups B.

Methods

POISSON–BOLTZMANN, STOCHASTIC DYNAMICS

The PBSD algorithm has been described elsewhere.[8,10] Briefly, the energy of the solvated system is separated into two components. The first, corresponding to the system in vacuum, is evaluated using a standard molecular mechanics potential with a uniform dielectric of 1. The second component represents the work associated with solvation of the rigid molecule. This term is further divided into the purely nonpolar work involved in the solvation of the molecule in a hypothetical neutral state, and the work required for the charging of the molecule from neutral to fully charged, in solution. The nonpolar term is calculated as the product of the molecule's solvent accessible surface area and a proportionality constant determined from studies of alkane solvation: 6 cal/Å^2.[23] The work associated with charging is obtained from solution of the Poisson–Boltzmann (PB) equation. In the PB model, the molecule is represented as a collection of point charges embedded in a low dielectric ($\epsilon = 1.0$) cavity which is, in turn, immersed in high dielectric ($\epsilon = 80.0$) solvent. This solvent can include ions represented in the PB equation through a bulk ionic strength. In the simulations reported here, the ionic strength was set to zero. The extent of the low dielectric cavity is determined by the van der Waals (vdw) radii of the atoms. This surface definition was used rather than the more common molecular surface[24] because the calculation of reaction field forces can fail due to numerical problems associated with the sudden appearance or disappearance of cavities and concavities in the surface.[8] These events can occur frequently if the molecular surface definition is used. Previous work has shown that reasonable results are obtained when using the vdw surface.[10] A radius of 1.0 Å was assigned to polar hydrogens due to numerical problems that arise in the solution of the PB equation when solvent exposed atoms of zero radius are used. The PB equation, or in this case simply the Poisson equation, was solved with finite differences on a rectangular grid with dimensions of $30 \times 30 \times 30$ using the UHBD software package.[25] The inter-

node spacing was varied to ensure that the longest dimension of the molecule occupied 0.7 of the length of the grid. This resulted in the spacing varying between approximately 0.8 and 1.0 Å. Previous work has shown that grid spacings as coarse as this, which may not, in general, yield accurate solvation energies for individual structures, are adequate for stochastic dynamics simulations.[10] This is apparently due to the fact that the systematic error resulting from the use of a coarse grid is small, and the random error is small relative to the stochastic forces utilized to model solvent buffeting of the solute. The boundary potentials of the grid were calculated as the sum of the dielectrically ($\epsilon = 80.0$) screened Coulomb potential of each atom in the molecule.

Forces were obtained from the gradients of the appropriate energy terms. The gradient of the surface area of the molecule was calculated using the method of Sridharan et al.[26] The forces associated with the solvent reaction field were calculated from the PB solution using the method of Gilson et al.[8] It is worth noting that the reaction field forces include not only the forces resulting from the interaction of the atomic charges with the solvent reaction field, but also the pressure exerted by the dielectric boundary on surface atoms of the molecule.

Because solution of the PB equation is time consuming, it is not practical to update the reaction field forces at every step of the simulation. Instead, the reaction field was recalculated every time an atom was found to have moved more than 0.3 Å from its position during the last calculation. This resulted in calculations occurring approximately every 8 to 12 fs in both simulations. Because calculation of surface area and its derivatives can also be resource intensive, and because surface area does not change rapidly, it was only updated when the reaction field was recalculated.

Because the PB solvent model does not include the effects of random buffeting and viscous drag exerted by the solvent, it is necessary to add them to ensure that a Boltzmann-weighted distribution of conformations is sampled. This can be accomplished through the use of a stochastic dynamics integrator.[14] At each step an additional random force, chosen from a Gaussian distribution of appropriate width, is added to each atom. A frictional force proportional to the atom's velocity is also added. The proportionality constant, referred to here as the coefficient of friction γ, together with the mean square width of the Gaussian stochastic force distribution, determines the system's temperature. In effect, the addition of stochastic forces acts to couple the explicit atoms to a thermal bath of a given temperature. A temperature of 300 K was used for the cryptophane simulations. The friction coefficient can be estimated from the bulk viscosity of water to be approximately 50 ps^{-1}. However, as suggested by previous reports, lowering γ can increase the rate at which conformational transitions occur.[16] Accordingly, a γ value of 10 ps^{-1} was used to accelerate sampling. While this change will affect the dynamics of the system, it should not affect equilibrium-average conformational properties.

Initial configurations for the PBSD simulations were taken from the structures at 1 ns in the previous explicit–solvent simulations. All water molecules were removed and velocities assigned from a Maxwellian distribution consistent with 300 K. Equilibration (100 ps) was performed to allow the system to relax in the continuum solvent environment. A 5-ns simulation was then carried out starting from the structure at the end of the equilibration.

PARAMETERS

To facilitate comparison of the continuum– and explicit–solvent simulations, an effort has been made to utilize atomic and molecular dynamics parameters consistent with the previously reported molecular dynamics simulations. A brief overview of the parameterization used is presented here. Details can be found in the reports describing the explicit simulations.[17–19]

To maintain adequate solubility in aqueous experimental studies, it is necessary to titrate the cryptophane molecule until four of its six carboxylic acid groups are ionized.[22] Therefore, in this, as in the previous explicit simulations, the cryptophane was modeled with four charged (ACET) and two neutral (ACID) acid groups. The chosen arrangement of groups is shown in Figure 1.

An all-atom representation was employed for the cryptophane, tetramethylammonium, and neopentane. Internal bonded parameters were taken from the AMBER force field.[27] Nonbonded parameters were obtained from two sources. Lennard–Jones parameters were taken from the OPLS parameter set with all-atom extensions.[28,29] All 1–4 interactions were scaled by $\frac{1}{8}$, except for polar–polar 1–4 interactions, which were scaled by $\frac{1}{2}$.

Polar–polar interactions were scaled in this way because it has been observed[30] that, in molecular dynamics simulations, the absence of vdw radii on polar hydrogens in the OPLS parameter set (originally designed for Monte Carlo simulations using rigid bond lengths and angles) can result in the formation of unrealistically stable five-membered, hydrogen-bonded, rings. In the cryptophane system this occurs due to a hydrogen bond between the ACID proton and the ether oxygen, which links the group to the rest of the cryptophane. Modified 1–4 scaling was suggested as a possible solution. By scaling the 1–4 polar interaction by $\frac{1}{2}$ rather than $\frac{1}{8}$, the repulsion is greater between the oxygens of the ether and ACID groups, destabilizing the ring conformation. Despite this, such rings formed frequently in both the explicit–solvent and PBSD simulations.

A RHF/6-31 + G* quantum mechanical calculation (using β-methoxyacetic acid as a model of the ether-linked ACID groups of the cryptophane) of the torsional potential of the acid hydrogen predicts a significant gas phase population of both rotamers (estimated ΔG of 0.6 kcal/mol in favor of the *cis* (nonring) conformation. However, the greater dipole moment of the *trans* conformer (5.7 vs. 2.1 D) could result in it being preferred in solution. This would be consistent with the results produced by both continuum and explicit solvent simulations using the current parameters.

The current parameterization, however, results in very close approaches between the polar ACID hydrogen and the ether oxygen (a minimum separation of 1.2 Å was observed). At such small separations their interaction can be so strong as to cause numerical problems in both explicit–solvent and continuum simulations if bond length constraints are not used. Simulations in which a reasonable radius was assigned to the polar hydrogen did not exhibit such instability (data not shown). However, to allow unambiguous comparison to the explicit–solvent simulations, no modifications were made to improve the quality of these interactions. Utilization of bond length constraints eliminated most of the problems. Extremely brief (a few 2 fs time steps) spikes of several hundred degrees in the temperature were observed. These were associated with angle stress in the five-membered rings, and occurred several times per ns of simulation. The extremely short duration of the events, and their consequently insignificant impact on the trajectory, allowed the simulations reported here to be carried out using the standard parameterization of OPLS (with modified 1–4 scaling).

The Coulombic parameters were obtained from *ab initio* molecular orbital calculations. As explained previously,[17] the partial charges were obtained by fitting atom-centered charges to the electrostatic potential of individual fragments of the cryptophane, and to the whole molecules for tetramethylammonium and neopentane.

All bond lengths were constrained using the SHAKE algorithm.[31] This allowed an integration timestep of 2 fs to be used. The distribution of dihedral angles in the cryptophane was monitored at every step (see below for details). Energies and coordinates were recorded at 0.05-ps intervals for further analysis.

EXPLICIT–SOLVENT SIMULATION DETAILS

For completeness, a few details concerning the previously reported 25-ns explicit–solvent simulations are recorded here. As noted in the last section, the atomic parameters were the same for both the explicit–solvent and PBSD simulations. The simulations were conducted at 298 K in the NPT ensemble using the molecular dynamics program ARGOS 6.0.[32] A 14.0-Å long-range nonbonded cutoff was used. The time step was 2 fs. Dihedral distributions were monitored at every step, with coordinates being recorded at 0.05-ps intervals for the empty host, and at 0.1 ps intervals for the TMA and NEO complexes.

Results and Discussion

STRUCTURAL CHARACTERISTICS

In the following discussion the relative length of the simulations should be kept in mind. The explicit–solvent simulations were run for five times as long as those in continuum–solvent.

The general structural behavior of the empty cryptophane was quite different from that observed in the explicit–solvent simulation. Whereas the cryptophane remained roughly spherical with ACID and ACET groups extended in explicit solvent, it sampled significantly different conformations in the PBSD simulation. Figure 2 contains a representative structure taken from the PBSD simulation. The presence of one of the ACID groups within the binding cavity can be seen. In fact, at approximately 2 ns the ACID group actually emerges from the cavity through a pore flanked by two ACET groups. This results in significant distortion of the cryptophane away from a spherical structure.

However, as was the case in the explicit–solvent simulations, the cryptophane remains roughly spherical throughout the 5 ns of both the cryptophane/TMA and cryptophane/NEO PBSD simulations. In both cases the ligand remained near the center of the binding cavity throughout the simulation. Representative structures from these simulations are also given in Figure 2.

The distribution of the radius of gyration of the cryptophane is plotted for all simulations in Figure 3 (three PBSD and three explicit simulations). In both continuum and explicit simulations the binding of a ligand slightly reduces the radius of gyration, with TMA causing the greatest reduction in explicit solvent. In the PBSD simulations, the TMA and NEO distributions are nearly identical and are similar to that of the explicit–solvent TMA simulation. The PBSD distribution for the empty cryptophane indicates a slightly more compact structure than is observed in the explicit–solvent simulation.

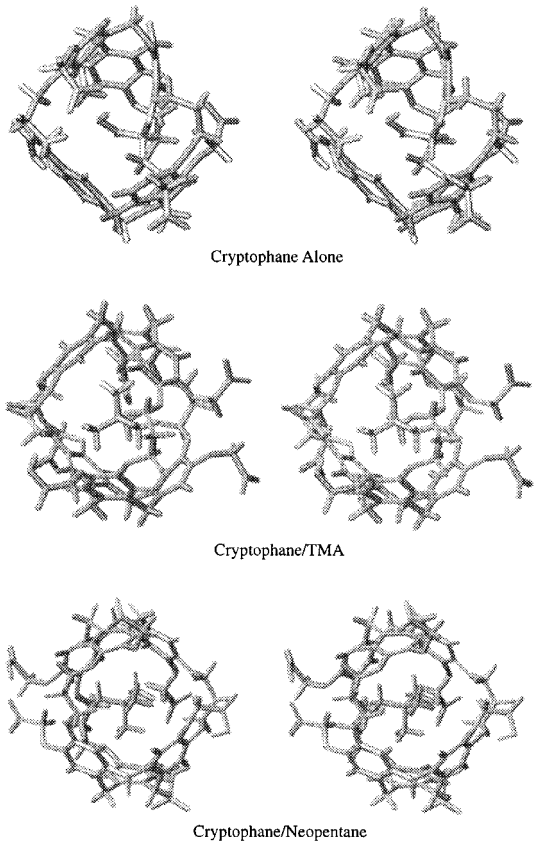

FIGURE 2. Structures taken from the cryptophane, cryptophane/TMA, and cryptophane/neopentane PBSD simulations. They are representative of the general structural characteristics observed throughout the simulations.

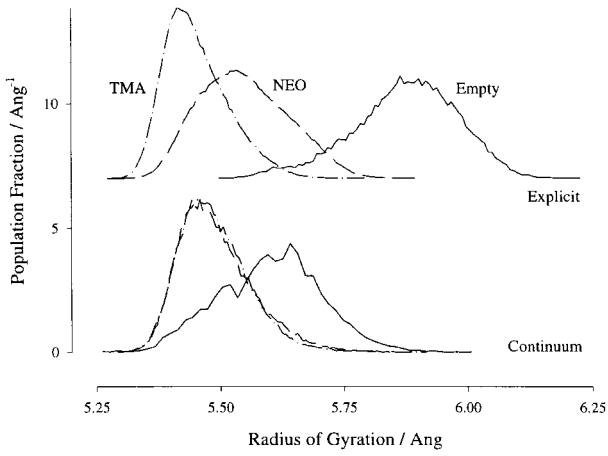

FIGURE 3. The distribution of the cryptophane radius of gyration for the PBSD and explicit–solvent simulations. The explicit–solvent data are offset vertically by +7.0.

Histograms of the ACET and ACID group distances to the center of geometry of the cryptophane are given for the three PBSD simulations in Figure 4 (top). The density present for the ACID groups at very small separation in the empty host simulation is due to autobinding of the ACID groups within the cryptophane. This does not occur when a ligand is present. This behavior could be evidence for inaccuracy in the continuum–solvent model; it is also possible that it results from enhanced sampling. For similar motion to occur in the explicit simulation, several water molecules would have to be dislodged from within the cavity. This would significantly slow down the process and might prevent it from occurring in 25 ns of simulation. However, one would expect a significant entropic penalty associated with the ''acid-in'' conformations. Thus, their prevalence in the empty continuum simulation does seem to indicate a problem with the solvent model.

A trend can be observed as one moves from the empty, to NEO, and then finally to the TMA simulation that is similar to that observed in the explicit data (Fig. 4, bottom). The ACID groups move outward and the ACET groups (10 and 13, see Fig. 1) move inward. The mobility of all groups is, restricted in the TMA and NEO simulations relative to the empty simulation. Although the inward movement of the two ACETs and outward shift of the ACIDs might be expected for the TMA complex, due to the charge on the ligand, the similarity of the TMA and NEO distributions indicates that the changes are a superimposition of the steric and electrostatic effects of ligand binding. In the

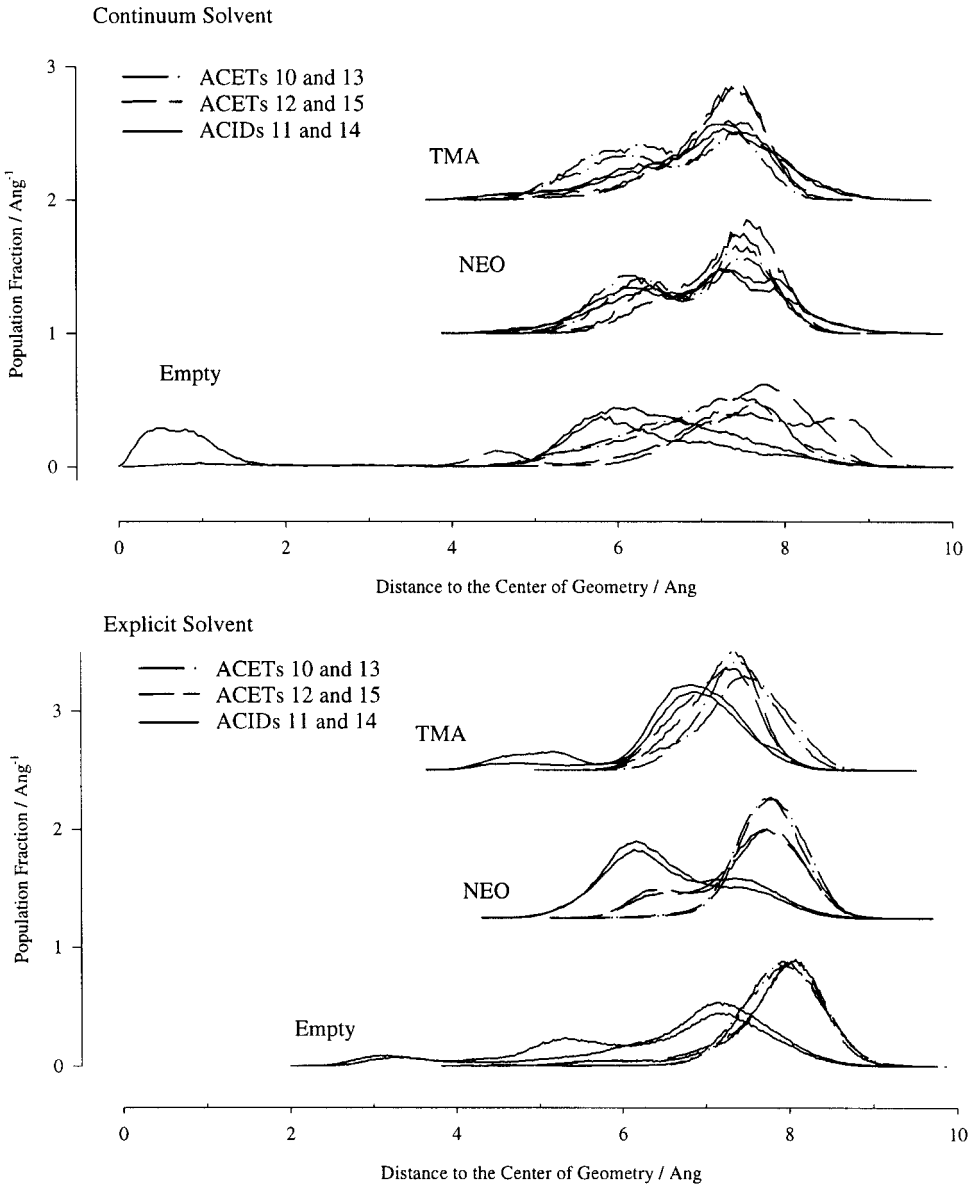

FIGURE 4. The distributions of ACET and ACID distances to the center of geometry of the cryptophane for the three PBSD (top) and three explicit solvent (bottom) simulations. Distances are measured from the carboxyl carbon of each group. The cryptophane/neopentane and cryptophane/TMA data are offset vertically to ease comparison.

explicit simulations, the charged ACET groups are significantly less mobile than the uncharged ACIDs. Greater ACET mobility in continuum solvent is expected given the strong interactions with solvent molecules that are present for charged groups in explicit solvent.

A comparison of the ACET and ACID group distances to the cryptophane center of geometry for the PBSD and explicit–solvent NEO and TMA simulations indicates that the main difference lies in the behavior of ACET groups 10 and 13. These are seen to have significant density at smaller distances in the PBSD simulations. Examination of the TMA trajectory indicates that these peaks are composed of ACET groups interacting closely with the TMA ligand. In both the TMA and NEO simulations these ACETs also share a pore with ACID groups to which they often hydrogen bond. These ligand and hydrogen bonding interactions (see below) occur much more frequently in continuum solvent, and are likely the cause of the differences in the distributions.

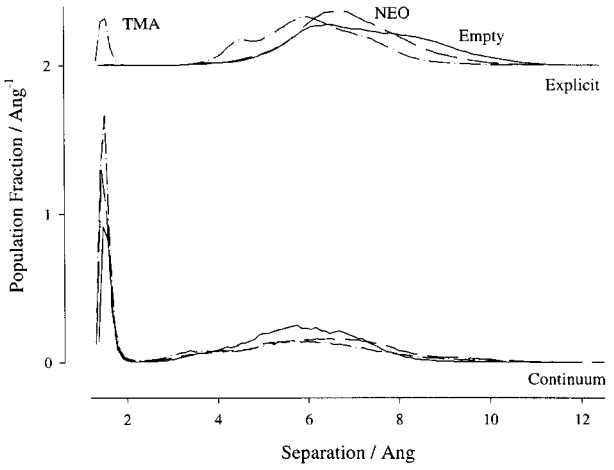

FIGURE 5. The distributions of the smallest separation between an ACET oxygen and the hydrogen of one of the ACID groups for the PBSD and explicit–solvent simulations. The explicit–solvent data are offset vertically by +2.0.

It is expected that differences in the solvent models will be most evident in the behavior of charged groups exposed to the solvent. Accordingly, the interaction of the ACET groups with the ACIDs and with each other were analyzed. Figure 5 is a comparison of the distribution of the shortest distance between one of the ACID protons and any ACET oxygen for explicit and continuum solvents. Very close ACID/ACET interactions are present for a significant portion of the PBSD simulations. There is an increase in the amount of pairing observed as one moves from the empty host, to NEO, to the TMA simulation. It is not clear why this should be the case. It is possible that the slight shift of the ACET groups closer to the center of geometry of the cryptophane (see Fig. 4) is associated with an increase in ACET/ACID encounters. The explicit simulation exhibits significant ACID/ACET pair formation only for the TMA simulation. The ''nonpaired'' portions of the distributions (greater than 2 Å separation) are shifted to slightly larger separations relative to the PBSD results. The corresponding distributions for the other acid proton are very similar in all cases except that of the TMA explicit simulation, in which no significant pairing was observed. Because the acid groups are equivalent, this difference indicates inadequacy in the sampling of these groups in the explicit simulation. The lack of equivalent sampling for the ACID groups is also evident in Figure 4. It is interesting to note that while the paired portion of the distribution (from approximately 1.3 to 2 Å separation) originates from contacts throughout the entire 5 ns for the PBSD simulation, pairs being formed and broken frequently, in explicit solvent, the density is almost entirely the result of a two periods of continuous contact of approximately 1 and 2 ns duration.

Figure 6 contains a comparison of the distribution of ACET/ACET separations for the six possible pairs for PBSD (top) and explicit (bottom) simulations. The distributions for the three simulations are similar in both continuum and explicit solvent. The distributions for the PBSD simulation are shifted very slightly towards smaller separations, relative to those in explicit solvent, for all but the pair at closest separation (12 of 15). This may be a reflection of the generally more compact structures sampled by the PBSD simulations (as indicated by Fig. 3). It is also possible that the apparent shift is contributed to by use of electrostatic cutoffs in the explicit simulations.

The abundance of ACET/ACID pairs observed in continuum solvent could indicate that the continuum–solvent reaction field effects are weaker than those found in explicit simulation. Reaction field forces operate in a direction opposite to Coulomb pair interactions. Solvation favors increase in charge density, and consequently, encourages association of like charges. Electrostatic potentials of mean force for the association of ions and other charged groups are thus determined by a sum of opposing reaction field and Coulombic components. Because these terms are often quite large, results can be very sensitive to errors in either component. It should be emphasized that when the continuum reaction field is referred to as ''weak,'' what is meant is that the reaction field component of the force acting on an atom is weaker in continuum solvent than it is in explicit solvent. Because the Coulomb evaluation is essentially identical for the PBSD and explicit–solvent simulations, differences in the behavior of charged groups can be attributed to differences in solvation arising from the different solvent models. One of the goals for this comparison of PBSD to explicit–solvent simulation, is to determine whether the continuum reaction field can be calculated with sufficient accuracy using atomic radii and charges which were optimized for explicit, rather than continuum solvent. While optimization of parameters specifically for continuum calculations is certainly desirable, it is common to utilize explicit–solvent parameterizations. This approach is convenient given the difficulty of generating transferable continuum parameters which reproduce both reaction

field forces and solvation energies.[33] The disparity in ACET/ACID pair formation in the PBSD and explicit simulations could indicate a problem with the use of unoptimized parameters. There are a number of possible explanations for this behavior. Two possibilities are a general weakness in the reaction field forces relative to those in explicit solvent, and a local problem caused by the use of 1.0 Å radii for polar hydrogens in the continuum electrostatic calculation, while they possess zero radius for vdw energy and force evaluation. Analysis of ACET/ACET pair separation, which one might expect to be further separated in the PBSD simulation if the reaction field was weak, does not indicate global weakness in the continuum reaction field. Most ACET/ACET pairs are actually slightly closer together in the PBSD simulations. It is interesting to note that the one pair that is shifted to larger separation is also the pair at closest separation (the ACETs that share a pore). It is this pair that one would expect to most strongly reflect a weakness in the continuum reaction field. All the shifts are, however, very small and may be a sampling artifacts. Additional simulations in

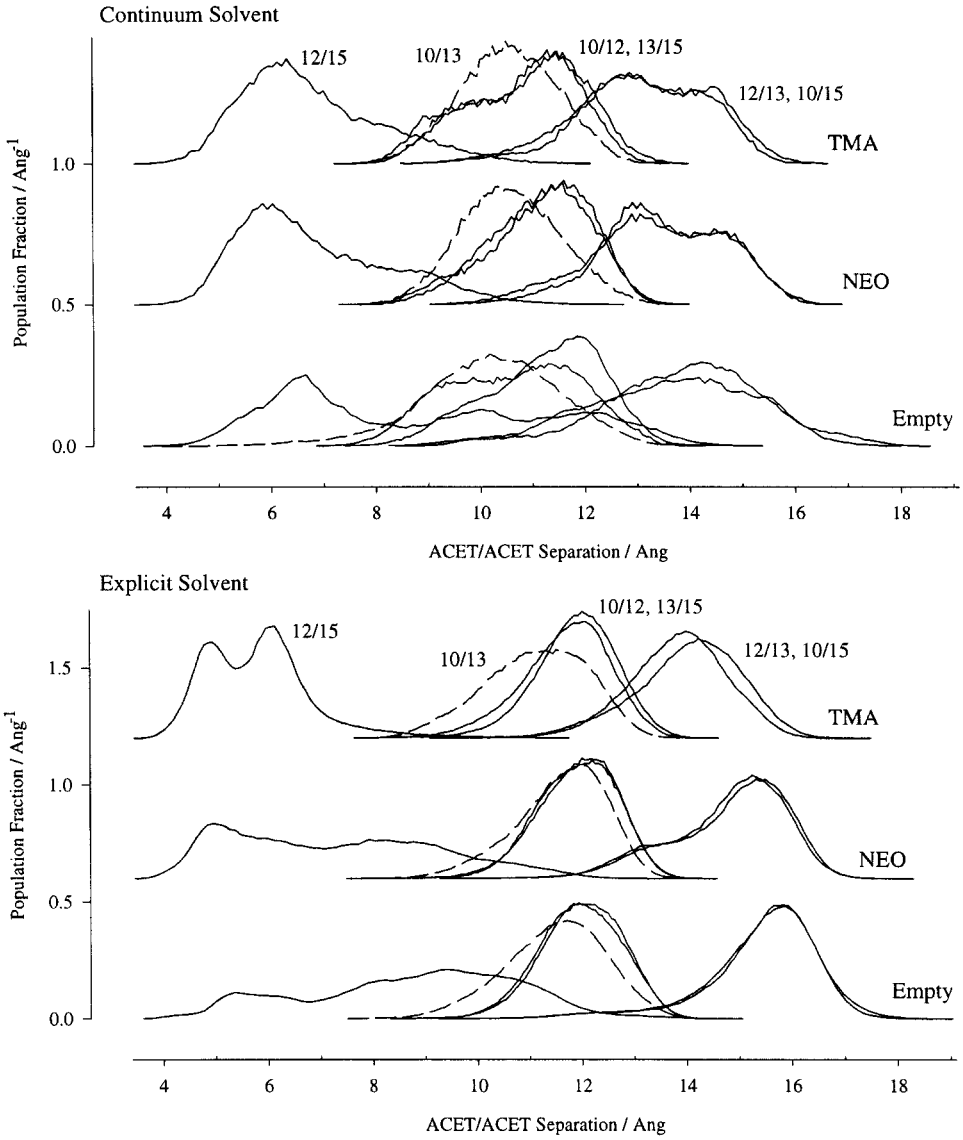

FIGURE 6. The distributions of the distances between ACET group carboxyl carbons for the empty host, cryptophane/NEO, and cryptophane/TMA PBSD (top) and explicit–solvent (bottom) simulations. Pairs are labeled with the group numbers of the participating ACETs (see Fig. 1). The cryptophane/NEO and cryptophane/TMA data are offset vertically to ease comparison.

which the polar hydrogen radii were made consistent for both electrostatic and vdw calculations show the same ACID/ACET behavior. So the hydrogen radii inconsistency does not give rise to the pairing behavior.

Enhanced sampling of ACET conformations in the PBSD simulations relative to explicit solvent (see the discussion of sampling rate below) provides an alternative explanation for the observed ACET/ACID pairing. The ACET/ACID pairing observed in the TMA explicit–solvent simulation supports this hypothesis. The fact that strong and persistent pairing is observed for a contiguous portion of this simulation indicates that pair formation is energetically reasonable in explicit solvent and that sampling differences could be responsible for the greater ACET/ACID pairing observed in the PBSD simulations.

ORIENTATION OF THE LIGAND

The orientation of the ligand within the binding cavity of the cryptophane molecule was characterized in explicit solvent[19] through analysis of the orientation of the ligand's four C3 rotational axes with respect to the cryptophane's C3 and three C2 axes (see Fig. 7). This analysis was carried out for the PBSD simulations. Details of the procedure can be found in a previous report,[19] and only a brief description is given here.

The four C3 axes of the ligands were defined as unit vectors oriented along the carbon–carbon and carbon–nitrogen bonds of neopentane and TMA, respectively. The C3 axis of the cryptophane was defined as a unit vector running between the centers of geometry of the two ''cap'' regions of the cryptophane. Only atoms contained in the ''caps'' were included in the calculation of the center of geometry, i.e., atoms in the linker segments and the carboxylic acid chains were not included (a precise definition of the ARYL fragments composing the cap region can be found in a previous report[17]). The three C2 axes of the cryptophane were defined as unit vectors passing through the center of geometry of the cryptophane and the center atoms of the three linker chains. The cryptophane center of geometry was calculated without consideration of the atoms comprising the carboxylic acid chains (the atoms used were those contained in the ARYL and PRPL fragments as defined previously[17]).

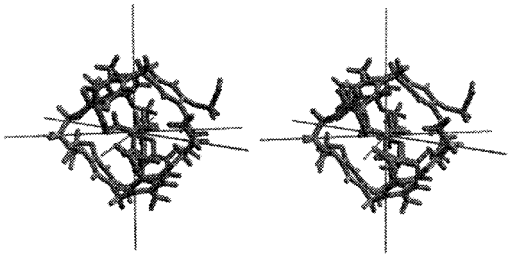

FIGURE 7. Illustration of the orientation and position of the cryptophane C3 (vertical) and three C2 axes (horizontal) used in the analysis of ligand orientation within the binding cavity.

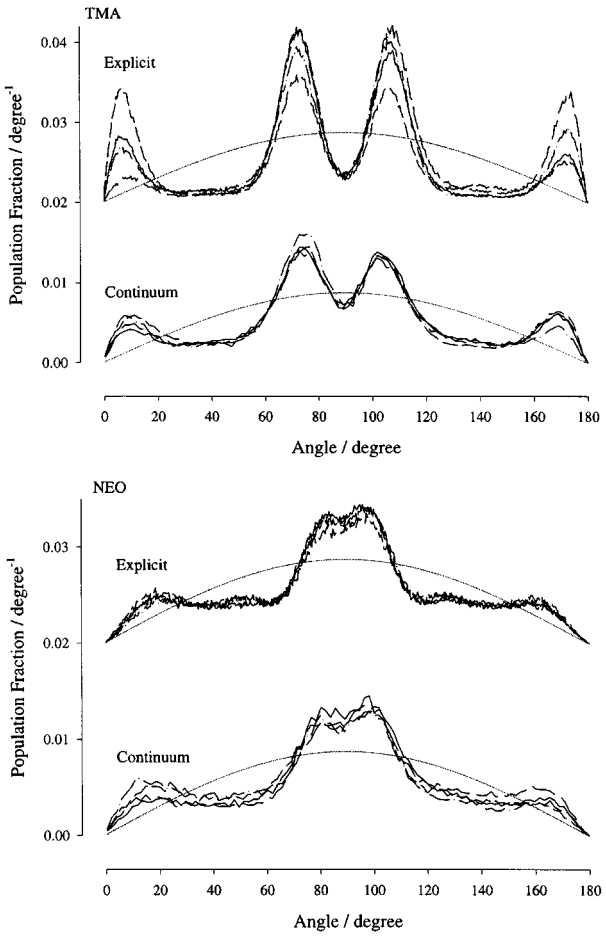

FIGURE 8. Distributions of the orientation of TMA's (top) and NEO's (bottom) C3 axes relative to the cryptophane C3 axis. Data are given for the PBSD and explicit–solvent simulations. The expected distributions for random ligand orientation are indicated by light gray lines. The explicit–solvent data are offset vertically by +0.02.

Figure 8 contains the distributions of the angles between the cryptophane C3 axis and the four C3 axes of TMA (top) and neopentane (bottom). For comparison, the distribution expected for completely random orientation is included as well.

This distribution is not uniform, due to the different effective sizes of the bins. This can be understood by visualizing the bins as sections of the surface of a sphere. The surface of a section at an angle of 90 degrees is larger than that at an angle close to 0 or 180. Figure 9 places the locations of the distribution peaks in the structural context of the cryptophane to aid in their interpretation. It is clear that, in the case of TMA, one of the methyl groups is generally oriented towards the center of one of the cap regions, with the three remaining groups constrained by the tetrahedral structure of TMA to fall at approximately 75 or 105 degrees (depending on which cap the first methyl group is pointed towards). For both ligands, the PBSD and explicit–solvent distributions are quite similar with respect to peak locations and widths. However, the explicit solvent distributions for the four axes of TMA (Fig. 8), are not equivalent, as would be expected. This indicates that there was insufficient sampling of the rotation of TMA about the cryptophane C3 axis even in 25 ns of explicit–solvent simulation. In contrast, the distributions from the PBSD simulation exhibit greater similarity. This indicates that better sampling was obtained with the continuum solvent model. The explicit–solvent and PBSD distributions also differ in the regions between the peaks. There is more density in these locations in the PBSD data. This indicates greater rotational freedom of the ligand, which must contribute to the better sampling. This additional freedom may be partially the result of the presence of high dielectric within the binding cavity surrounding TMA. Because the vdw surface of the atoms is used to determine the high and low dielectric regions for the continuum electrostatics calculation, any region not overlapped by an atom is assigned a high dielectric. This occurs regardless of whether there is actually enough space at the location to accommodate a water molecule. This could reduce the strength of electrostatic interactions between the ligand the host and, together with the effectively instantaneous relaxation that occurs on each PB update, result in considerably greater rotational freedom than is observed in the explicit–solvent simulation.

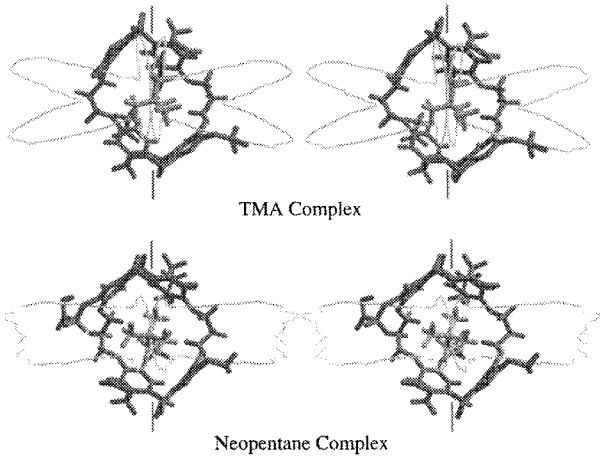

FIGURE 9. Distributions of the orientation of one of the ligand C3 axes relative to the cryptophane C3 axis. The cryptophane C3 is oriented vertically, and passes through the center of the cryptophane in each figure (its orientation is indicated by small black lines). Data are given for the PBSD cryptophane/TMA and cryptophane/neopentane complexes. The data are mapped to a circle with the radius proportional to the distribution value. A portion of the cryptophane has been removed to improve the visibility of the distribution.

The distributions for the explicit solvent and PBSD neopentane simulations are nearly identical (see Fig. 8). As was found in the explicit simulation, the neopentane distribution has much broader peaks and is relatively flat overall, particularly when compared to the random distribution. It is clear that neopentane is not as rotationally constrained as TMA. This lack of constraint results in much better sampling as evidenced by the similarity of the distributions for the four C3 axes of the ligand.

The angle between the ligand C3 axes and that of the cryptophane only provides information concerning the ligand's orientation about the cryptophane C3 axis. Additional information can be obtained from the angles between the ligand C3 and cryptophane C2 axes. Figure 10 contains the distributions of these angles for the TMA (top) and NEO (bottom) simulations. The distributions for the four equivalent ligand C3 axes have been averaged together for each cryptophane C2 to ease presentation of the data. Because of the protonation pattern assumed for the cryptophane, two of the three resulting distributions should be equivalent (the two C2 axes that emerge from pores containing one ACET and one ACID group). The explicit solvent and PBSD distributions for the TMA simulations are quite similar. As was the case above, the PBSD peaks are slightly broader, indicating the slightly greater rotational freedom of the ligand in continuum solvent. The locations of the peaks are consistent with one methyl group being pointed up into the cap region and the others twisted approximately 20 degrees away from perfect alignment with the pores of the cryptophane. Orientation of the groups towards the

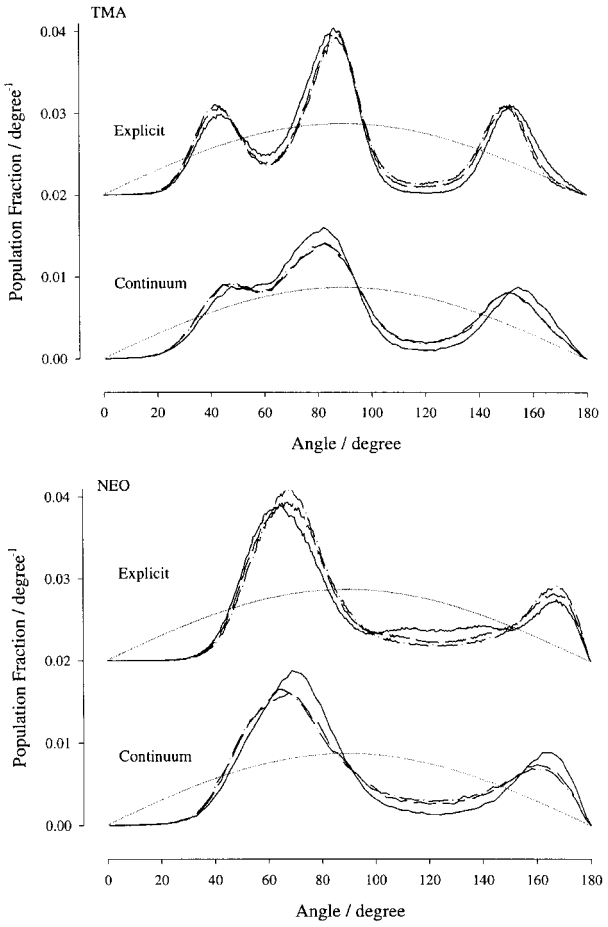

FIGURE 10. Distributions of the orientation of the ligand C3 axes relative to the cryptophane C2 axes. The contributions from the four ligand C3 axes were averaged to yield one curve relative to each cryptophane C2 (for a total of three). Data are given for the PBSD and explicit–solvent cryptophane/TMA (top) and NEO (bottom) complex simulations. The expected distributions for random ligand orientation are indicated by light gray lines. The explicit–solvent data are offset by +0.02.

pores presumably occurs for steric reasons. The twist may be the result of electrostatic interactions between the methyl hydrogens of TMA and the ether oxygens of the chains that connect the cryptophane caps.

The explicit–solvent and PBSD distributions for the neopentane simulations are quite similar to each other. They are consistent with the orientation of one methyl group towards a cap region, and the others towards the centers of the pores. However, there is apparently considerable freedom for rotation relative to the TMA complex. This results in the blending together of what might ideally be peaks at 60 and 90 degrees into a single broad peak. This is consistent with the greater rotational freedom about the cryptophane C3 axis observed for the neopentane complex.

SAMPLING RATE

In the discussion above, it was observed that PBSD simulation often produces enhanced sampling relative to that observed for explicit solvent simulations. This is likely due to a combination of the use of an artificially low friction coefficient and the properties of the continuum solvent model (instantaneous solvent relaxation, presence of high dielectric inside the binding cavity, etc.). To quantify any advantages in sampling rate provided by the continuum solvent model, the rate at which new unique conformations were discovered was computed by monitoring the cryptophane's dihedral angles. This was accomplished in the same manner as was done for the previous explicit and PBSD simulations.[17] Each angle was divided into wells. Divisions between wells were placed at the maxima in the dihedral potential. At each molecular dynamics step, every angle was assigned an integer corresponding to the well currently occupied by the angle. This resulted in a string of integers characterizing the conformation. A conformation was considered a new unique structure if an equivalent string (equal or equivalent by symmetry) had not yet been recorded. Angles that involved hydrogens were not considered in the analysis.

For all simulations, the sampling rate is essentially constant throughout the run, indicating that sampling was not exhaustive. The PBSD simulations sample at a higher rate in all simulations, with the largest advantage occurring in the TMA simulation where more than three times as many conformations are sampled (based on a comparison of 5 ns of PBSD to 5 ns of explicit–solvent simulation).

This initial conformational analysis was further refined by dividing the cryptophane into ''exo'' and ''cage'' regions. The exo region was defined as consisting of the ACID and ACET groups, including the ether oxygens. The remainder, which forms the core of the molecule and has the largest role in defining the size and shape of the binding cavity, was assigned to the cage region.

Sampling of the cage region is greatly restricted by the presence of a ligand in both the PBSD and explicit–solvent simulations. This restriction is most severe for the TMA complex. While approximately 20000 unique cage conformations where

sampled in the PBSD empty simulation, only 1200 where found with TMA bound. A similar degree of conformational restriction is observed in the explicit data. The continuum solvent simulation does sample the cage region of the TMA complex at a much higher rate than is observed for the explicit simulation, locating more than eight times as many unique conformations in 5 ns. It is clear from the rotational analysis above that TMA exhibits considerably greater rotational freedom within the cryptophane in continuum solvent. This, combined with the much faster sampling of the cage region, is illustrative of the enhanced sampling of the continuum solvent model. In contrast to the TMA simulation, the PBSD and explicit–solvent sampling rates are nearly identical for the cage region of the neopentane complex. This is what one might expect if the interactions between neopentane and the cryptophane are largely the result of vdw contacts that are treated identically in the PBSD and explicit–solvent simulations.

For all simulations PBSD also samples the exo region at a higher rate, generating more than three times as many unique structures as the explicit–solvent simulations. This increase in the rate of sampling of the exo region, which consists of carboxylic acid groups with their associated linker chains, indicates the sampling advantages of the continuum solvent model for charged groups which interact strongly with solvent.

Table I summarizes the results of the sampling analysis. The sampling enhancement resulting from the use of the continuum solvent model increases as one moves from the empty host, to the neopentane complex, and finally to the TMA complex, for which the PBSD simulation sampled more than three times as many conformations as the explicit. The data for the cage and exo regions illustrates how nonuniform the sampling enhancement can be within a given system. In general, one sees the greatest enhancement for regions of the structure that interact strongly with solvent and require solvent displacement for conformational change.

TABLE I.
Ratio of the Number of Unique Conformations Located by PBSD to the Number Found by Explicit–Solvent-Simulation in 5 ns.

System	Whole	Cage	Exo
Empty host	1.2	1.8	3.0
Neopentane complex	1.5	0.92	3.3
TMA complex	3.2	8.2	3.3

The cage region of the TMA simulation was particularly well sampled in the continuum simulation, possibly due to weakening of the electrostatic interactions between TMA and the cryptophane due to the continuum dielectric representation.

Conclusion

The comparison of PBSD to explicit solvent simulation in the cryptophane system has provided important information regarding the advantages and disadvantages of the PBSD approach. One of the primary questions addressed by this study is the adequacy of the continuum solvent model when utilizing parameters designed for use in explicit solvent. Here, adequacy is defined as the ability to reproduce the results of explicit–solvent simulation. The apparent preference of the neutral ACID groups for the interior of the binding cavity in the PBSD empty cryptophane simulation is one important discrepancy. Another significant difference in the simulations that casts doubt on the quality of the model is the large number of close ACID/ACET interactions that are observed in all of the PBSD simulations, but found only in the TMA simulation in explicit solvent. This could result from a weakness in the PBSD reaction field relative to explicit solvent. However, the sampling advantages of PBSD for the ACID and ACET groups, combined with the formation of strong and persistent ACID/ACET pairs in the explicit–solvent TMA simulation complicates interpretation. It is possible that the lack of such pairs in the other explicit–solvent simulations may be due to inadequate sampling. Similarly, the requirement for solvent displacement in autobinding of the ACIDS could result in bound conformations never being observed in explicit simulation, regardless of their stability. Therefore, while these results do not conclusively show that the use of explicit solvent parameters leads to general qualitative inaccuracy, they do suggest that it is necessary that efforts be focused on development of continuum optimized parameters in cases where greater than qualitative accuracy is required. Until such parameters are available, the results presented here should at least provide an example of the magnitude and character of the differences one might expect between explicit–solvent simulation and PBSD.

In general, the consistency of other aspects of the PBSD results with those from explicit solvent is encouraging. Several qualitative conclusions

drawn from the explicit solvent data, concerning the equilibrium-average properties of the cryptophane and its complexes, can also be drawn from the PBSD data (effect of binding on conformational freedom, orientation of ligands, etc.).

Analysis of sampling rates shows that the PBSD method offers a large sampling enhancement relative to explicit solvent. It is also clear that the greatest enhancements are observed for portions of the system that interact strongly with solvent and that require solvent displacement for conformational transitions.

The PBSD simulations required approximately 200 cpu hours per ns on an Indigo2 R10000 SGI. Explicit–solvent simulation with ARGOS requires approximately 400 hours per ns. Although this is a significant increase in efficiency, the maximum computational benefits of PBSD with respect to cpu time/ns are not expected to be achieved in such a small system (160 atoms). The eventual goal is the application of PBSD to large macromolecular systems containing many thousands of atoms. The resource savings associated with a continuum solvent model are likely to be greater in such systems. It should also be noted that, due to the increased rate of sampling in PBSD, equivalent sampling may be possible to obtain in most systems with significantly smaller amounts of simulation. In the case of the cryptophane system, some properties, such as the rotational orientation of the ligand, were actually better sampled in 5 ns of PBSD simulation than in 25 ns of explicit. In addition, the PBSD software is not fully optimized. We are currently working to move from the current finite difference PB solver to one that utilizes multigrid methods[34] to yield a significant performance enhancement. It may also be possible to further simplify the continuum electrostatic calculation through the use of the generalized Born-solvation model.[35, 36]

Although PBSD is a promising method, further work in the area of parameterization and validation must be done in order for this method to become an efficient alternative to explicit simulation. Despite the problems noted above, it does produce reasonable qualitative results with current, explicit–solvent parameters in some cases. It also offers significant sampling advantages relative to explicit–solvent simulation, as well as a significant increase in computational efficiency. With further development, PBSD and related continuum–solvent methods should find wide application in the efficient simulation of solvated systems.

Acknowledgments

The authors would like to thank Dr. Adrian Elcock and other members of the McCammon group for their suggestions, comments, and support of this study. This work was supported in part by grants from the National Science Foundation (NSF) and the NSF Supercomputer Centers MetaCenter Program. M. J. P. was a Howard Hughes Predoctoral Fellow. H. A. C. thanks the American Cancer Society and the La Jolla Interfaces in Science Training Program.

References

1. Honig, B.; Sharp, K. A.; Yang, A. S. J Phys Chem 1993, 97, 1101.
2. Sharp, K. A. In Simulation of Biomolecular Systems: Theoretical and Experimental Applications; van Gunsteren, W. F.; Weiner, P. K.; Wilkinson, A. J., Eds.; ESCOM: Leiden, 1993, p. 147, vol. 2.
3. Sitkoff, D.; Sharp, K. A.; Honig, B. J Phys Chem 1994, 98, 1978.
4. Schmidt, A. B.; Fine, R. M. Mol Sim 1994, 13, 347.
5. Honig, B.; Nicholls, A. Science 1995, 268, 1144.
6. Sharp, K. J Comput Chem 1991, 12, 454.
7. Neidermeier, C.; Schulten, K. Mol Sim 1992, 8, 361.
8. Gilson, M. K.; Davis, M. E.; Luty, B. A.; McCammon, J. A. J Phys Chem 1993, 97, 3591.
9. Elcock, A. H.; McCammon, J. A. J Am Chem Soc 1996, 118, 3787.
10. Gilson, M. K.; McCammon, J. A.; Madura, J. D. J Comput Chem 1995, 16, 1081.
11. Smart, J. L.; Marrone, T. J.; McCammon, J. A. J Comput Chem 1997, 18, 1750.
12. Wan, S. Z.; Wang, C. X.; Xiang, Z. X.; Shi, Y. Y. J Comput Chem 1997, 18, 1440.
13. Wang, C. X.; Wan, S. Z.; Xiang, Z. X.; Shi, Y. Y. J Phys Chem 1997, 101, 230.
14. van Gunsteren, W. F.; Berendsen, H. J. C. Mol Sim 1988, 1, 173.
15. Elcock, A. H.; Potter, M. J.; McCammon, J. A. In Computer Simulation of Biomolecular Systems; van Gunsteren, W. F.; Weiner, P. K.; Wilkinson, A. J., Eds., ESCOM: Leiden, 1998, p. 244, vol. 3.
16. Loncharich, R. J.; Brooks, B. R.; Pastor, R. W. Biopolymers 1992, 32, 523.
17. Kirchhoff, P. D.; Bass, M. B.; Hanks, B. A.; Briggs, J. M.; Collet, A.; McCammon, J. A. J Am Chem Soc 1996, 118, 3237.
18. Kirchhoff, P. D.; Dutasta, J. P.; Collet, A.; McCammon, J. A. J Am Chem Soc 1997, 119, 8015.

19. Kirchhoff, P. D.; Dutasta, J. P.; Collet, A.; McCammon, J. A. J Am Chem Soc 1999, 121, 381.
20. Canceill, J.; Cesario, M.; Collet, A.; Guilhem, J.; Lacombe, L.; Lozach, B.; Pascard, C. Angew Chem Int Ed Engl 1989, 28, 1246.
21. Collet, A.; Dutasta, J. P.; Lozach, B. Adv Supramol Chem 1993, 3, 1.
22. Garel, L.; Lozach, B.; Dutasta, J. P.; Collet, A. J Am Chem Soc 1993, 115, 11652.
23. Simonson, T.; Brunger, A. T. J Phys Chem 1994, 98, 4683.
24. Richards, F. M. Annu Rev Biophys Bioeng 1977, 6, 151.
25. Davis, M. E.; Madura, J. D.; Luty, B. A.; McCammon, J. A. Comput Phys Commun 1991, 62, 187.
26. Sridharan, S.; Nicholls, A.; Sharp, K. A. J Comput Chem 1995, 16, 1038.
27. Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. J Comput Chem 1986, 7, 230.
28. Jorgensen, W. L.; Tirado–Rives, J. J Am Chem Soc 1988, 110, 1657.
29. Jorgensen, W. L.; Tirado–Rivers, J. J Am Chem Soc 1990, 112, 4768.
30. Smith, P. E.; Dang, L. X.; Pettitt, B. M. J Am Chem Soc 1991, 113, 67.
31. Ryckaert, J. P.; Cicotti, G.; Berendsen, H. J. C. J Comput Phys 1977 23, 327.
32. Straatsma, T. P.; McCammon, J. A. J Comput Chem 1990, 11, 943.
33. Jayaram, B.; Liu, Y.; Beveridge, D. L. J Chem Phys 1998, 109, 1465.
34. Holst, M. J.; Saied, F. J Comput Chem 1995, 16, 337.
35. Qiu, D.; Shenkin, P. S.; Hollinger, F. P.; Still, W. C. J Phys Chem A 1997, 101, 3005.
36. Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. J Am Chem Soc 1990, 112, 6127.